Journal of Computational and Applied Mathematics 123 (2000) 131–153

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

An overview of relative $\sin \Theta$ theorems for invariant subspaces of complex matrices ☆

Ilse C.F. Ipsen *

Center for Research In Scientific Computation, Department of Mathematics, North Carolina State University, P.O. Box 8205, Raleigh, NC 27695-8205, USA

Received 28 May 1999; received in revised form 15 September 1999

Abstract

Relative perturbation bounds for invariant subspaces of complex matrices are reviewed, with emphasis on bounding the sines of the largest principal angle between two subspaces, i.e. $\sin \Theta$ theorems. The goal is to provide intuition, as well as an idea for why the bounds hold and why they look the way they do. Relative bounds have the advantage of being better at exploiting structure in a perturbation than absolute bounds. Therefore the reaction of subspaces to relative perturbations can be different than to absolute perturbations. In particular, there are certain classes of relative perturbations to which subspaces of indefinite Hermitian matrices can be more sensitive than subspaces of definite matrices.

MSC: 15-02; 15A18; 15A42; 65F15; 65F35

Keywords: Invariant subspace; Eigenspace; Relative bound; Eigenvalues; Grading; Scaling

1. Introduction

The goal is to assess the quality of perturbed invariant subspaces of complex matrices. Of interest is a new class of perturbation bounds, called relative perturbation bounds. Relative bounds are better at exploiting structure in a perturbation than absolute bounds. In particular, relative bounds can be sharper than traditional bounds when the perturbations arise from numerical errors of certain computational methods. The following example illustrates what we mean by relative bounds.

☆ This research was supported in part by NSF grant DMS-9714811.

* Corresponding author.

E-mail address: ipsen@math.ncsu.edu (I.C.F. Ipsen); http://wwwy.ncsu.edu/˜ipsen/info.html

PII: S0377-0427(00)00404-0

Example 1.1 (*Ipsen* [18, *Example* 1]). Suppose

$$A \equiv \begin{pmatrix} a & & \\ & b & \\ & & c \end{pmatrix},$$

is a complex diagonal matrix of order 3 with distinct eigenvalues a, b, and c; and

$$A + E_1 \equiv \begin{pmatrix} a & \varepsilon & \varepsilon \\ & b & \varepsilon \\ & & c \end{pmatrix}$$

is a perturbed matrix with the same eigenvalues as A. We want to compare the eigenvectors of A and $A + E_1$ associated with eigenvalue c. The matrix A has $(0\ 0\ 1)^{\mathrm{T}}$ as an eigenvector [1] associated with c, while $A + E_1$ has

$$\begin{pmatrix} \frac{\varepsilon}{c-a}\left(\frac{\varepsilon}{c-b}+1\right) & \frac{\varepsilon}{c-b} & 1 \end{pmatrix}^{\mathrm{T}}.$$

The difference between these two eigenvectors depends on $\varepsilon/(c-a)$ and $\varepsilon/(c-b)$. This suggests that the angle between the two vectors can be bounded in terms of

$$\|E_1\|/\min\{|c-a|, |c-b|\}, \tag{1.1}$$

as $\|E_1\| = \mathcal{O}(|\varepsilon|)$.

Now consider the perturbed matrix

$$A + E_2 \equiv \begin{pmatrix} a & a\varepsilon & a\varepsilon \\ & b & b\varepsilon \\ & & c \end{pmatrix},$$

with the same eigenvalues as A. Again, compare eigenvectors of A and $A + E_2$ associated with eigenvalue c. An eigenvector of $A + E_2$ associated with eigenvalue c is

$$\begin{pmatrix} \frac{\varepsilon a}{c-a}\left(\frac{\varepsilon b}{c-b}+1\right) & \frac{\varepsilon b}{c-b} & 1 \end{pmatrix}.$$

The difference between the eigenvectors of A and $A+E_2$ depends on $\varepsilon a/(c-a)$ and $\varepsilon b/(c-b)$. This suggests that their angle can be bounded in terms of

$$\|A^{-1}E_2\|/\min\left\{\frac{|c-a|}{|a|}, \frac{|c-b|}{|b|}\right\}, \tag{1.2}$$

as $\|A^{-1}E_2\| = \mathcal{O}(|\varepsilon|)$.

Bound (1.1) is a traditional, absolute bound and $\min\{|c-a|, |c-b|\}$ is an absolute eigenvalue separation, while (1.2) is a relative bound and $\min\{|c-a|/|a|, |c-b|/|b|\}$ is a relative eigenvalue separation. □

The absolute bound contains $\|E\|$ and an absolute separation, while the relative bound contains $\|A^{-1}E\|$ and a relative separation. This means, the absolute bound measures sensitivity with regard to perturbations E, while the relative bound measures sensitivity with regard to perturbations $A^{-1}E$.

[1] The superscript T denotes the transpose.

The sensitivity to absolute perturbations is determined by an absolute separation, while the sensitivity to relative perturbations is determined by a relative separation.

There are other ways to construct relative bounds, by taking advantage of structure in the perturbation. The estimates provided by absolute and relative bounds can be very different. Which bound to use depends on the particular matrix and perturbation. One does not know yet in general which type of bound gives the best result for a given matrix and perturbation.

One advantage of relative perturbation bounds is that they can explain why some numerical methods are much more accurate than the traditional, absolute bounds would predict. That is because the errors caused by these methods can be expressed as small, relative perturbations. Specifically for the computation of eigenvectors, numerical methods that deliver high relative accuracy include:

- Inverse iteration for real symmetric scaled diagonally dominant matrices [1, Section 11] and real symmetric positive-definite matrices [8, Section 5].
- Two-sided Jacobi methods for real symmetric positive-definite matrices [8, Section 3].
- QR algorithms for real symmetric tridiagonal matrices with zero diagonal [6, Sections 5 and 6].
- Cholesky factorization followed by SVD of Cholesky factor for scaled diagonally dominant tridiagonals [1, Section 10] and for symmetric positive-definite matrices [7, Section 12]; [8, Section 4.3]; [23].
- Shifted Cholesky factorization followed by inverse iteration for real symmetric tridiagonal matrices [9, Section 5]; [10]; [28, Section 1].

Relative bounds are better at exploiting structure in perturbations than absolute bounds. For instance, from the point of view of absolute bounds there is no need to distinguish between definite and indefinite Hermitian matrices when it comes to sensitivity of invariant subspaces. However, from the point of view of relative bounds subspaces of indefinite Hermitian matrices can be more sensitive to perturbations than those of definite matrices for certain classes of perturbations, see Sections 3.3–3.6.

1.1. Overview

This article is a successor to the survey on relative perturbation bounds for eigenvalues and singular values [19] and a previous review [29]. Here we review relative perturbation bounds for invariant subspaces. Due to space limitations the emphasis is on bounding the sines of the largest principal angle between two subspaces, i.e. $\sin\Theta$ theorems. Some information can get lost by focussing on an angle. For instance, $\sin\Theta$ theorems give no information about the accuracy of individual eigenvector components. Such bounds on individual components are derived, for instance, in [1, Section 7]; [8, Section 2]; [25, Section 3]; [23, Theorem 3.3]; [22, Theorem 4].

The goal is to provide intuition, as well as an idea for why the bounds hold and why they look the way they do. We present and derive relative as well as absolute bounds to show that there is nothing inherently special about relative bounds. Sometimes relative bounds are even implied by absolute bounds, hence they are not necessarily stronger than absolute bounds.

Relative bounds have been derived in the context of two different perturbation models:

- *Additive* perturbations (Section 3) represent the perturbed matrix as $A+E$. Bounds for the following matrix types are presented: general (Section 3.1), diagonalizable (Section 3.2), Hermitian

positive-definite (Section 3.3), graded Hermitian positive-definite (Section 3.4), Hermitian indefinite (Section 3.5), and graded Hermitian indefinite (Section 3.6).
- *Multiplicative* perturbations (Section 4) represent the perturbed matrix as $D_1 A D_2$, where D_1 and D_2 are nonsingular matrices. Bounds are presented for diagonalizable (Section 4.1) and Hermitian matrices (Section 4.2).

1.2. Notation

Individual elements of a matrix A are denoted by a_{ij}. We use two norms: the two-norm

$$\|A\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} \quad \text{where } \|x\|_2 \equiv \sqrt{x^*x}$$

and the superscript $*$ denotes the conjugate transpose; and the Frobenius norm

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{i,j} |a_{ij}|^2}.$$

The norm $\|\cdot\|$ stands for both, Frobenius and two-norm. The identity matrix of order n is

$$I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = (e_1 \ldots e_n)$$

with columns e_i.

For a complex matrix Y, range(Y) denotes the column space, Y^{-1} is the inverse (if it exists) and $Y^\dagger$ the Moore–Penrose inverse. The two-norm condition number with respect to inversion is $\kappa(Y) \equiv \|Y\|_2 \|Y^\dagger\|_2$.

2. The problem

Let A be a complex square matrix. A subspace $\mathscr{S}$ is an *invariant subspace* of A if $Ax \in \mathscr{S}$ for every $x \in \mathscr{S}$ [15, Section 1.1]; [34, Section I.3.4]. Applications involving invariant subspaces are given in [15].

Let $\hat{\mathscr{S}}$ be a perturbed subspace. The distance between the exact space $\mathscr{S}$ and the perturbed space $\hat{\mathscr{S}}$ can be expressed in terms of $\|P\hat{P}\|$, where P is the orthogonal projector onto $\mathscr{S}^\perp$, the orthogonal complement of $\mathscr{S}$, while $\hat{P}$ is the orthogonal projector onto $\hat{\mathscr{S}}$ [18, Section 2]. When $\mathscr{S}$ and $\hat{\mathscr{S}}$ have the same dimension, the singular values of $P\hat{P}$ are the sines of the principal angles between $\mathscr{S}$ and $\hat{\mathscr{S}}$ [16, Section 12.4.3]; [34, Theorem I.5.5]. Therefore, we set

$$\sin\Theta \equiv P\hat{P}.$$

We present absolute and relative bounds for $\|\sin\Theta\|$, where $\|\cdot\|$ is the two-norm or the Frobenius norm.

3. Additive perturbations for invariant subspaces

The perturbed subspace $\mathscr{S}$ is interpreted as an exact subspace of a perturbed matrix $A+E$. Relative and absolute bounds on $\|\sin\Theta\|$ are presented for the following matrix types: general, diagonalizable, Hermitian positive-definite, graded Hermitian positive-definite, Hermitian indefinite and graded Hermitian indefinite.

3.1. General matrices

Absolute and relative bounds for invariant subspaces of complex square matrices are presented. The bounds make no reference to subspace bases and provide a unifying framework for subsequent bounds. They also illustrate that relative bounds exist under the most general of circumstances.

We start with the absolute bound. Define the absolute separation between A and $A+E$ with regard to the spaces $\mathscr{S}$ and $\hat{\mathscr{S}}$ by

$$\text{abssep} \equiv \text{abssep}_{\{A,A+E\}} \equiv \min_{\|Z\|=1, PZ\hat{P}=Z} \|PAZ - Z(A+E)\hat{P}\|,$$

where P is the orthogonal projector onto $\mathscr{S}^{\perp}$, and $\hat{P}$ is the orthogonal projector onto $\hat{\mathscr{S}}$. The absolute bound below holds for any square matrix.

Theorem 3.1 (Ipsen [18, Theorem 3.1]). *If* abssep > 0 *then*

$$\|\sin\Theta\| \leqslant \|E\|/\text{abssep}_{\{A,A+E\}}.$$

Proof. From $-E = A-(A+E)$ follows

$$-PE\hat{P} = PA\hat{P} - P(A+E)\hat{P}.$$

Since $\mathscr{S}^{\perp}$ is an invariant subspace of A^*, the associated projector P satisfies $PA = PAP$. Similarly, $(A+E)\hat{P} = \hat{P}(A+E)\hat{P}$. Hence

$$-PE\hat{P} = PA\sin\Theta - \sin\Theta(A+E)\hat{P}$$

and $\sin\Theta = P\sin\Theta\hat{P}$ implies

$$\|E\| \geqslant \|PE\hat{P}\| \geqslant \text{abssep}\|\sin\Theta\|. \qquad \square$$

Thus, the subspace $\mathscr{S}$ is insensitive to absolute perturbations E if the absolute separation is large.

Now we derive the corresponding relative bound. Define the relative separation between A and $A+E$ with regard to the spaces $\mathscr{S}$ and $\hat{\mathscr{S}}$ by

$$\text{relsep} \equiv \text{relsep}_{\{A,A+E\}} \equiv \min_{\|Z\|=1, PZ\hat{P}=Z} \|PA^{-1}(PAZ - Z(A+E)\hat{P})\|,$$

where P is the orthogonal projector onto $\mathscr{S}^{\perp}$, and $\hat{P}$ is the orthogonal projector onto $\hat{\mathscr{S}}$. The relative bound below holds for any nonsingular matrix.

Theorem 3.2 (Ipsen [18, Theorem 3.2]). *If A is nonsingular and* relsep > 0 *then*

$$\|\sin\Theta\| \leqslant \|A^{-1}E\|/\text{relsep}_{\{A,A+E\}}.$$

Proof. From $-A^{-1}E = I - A^{-1}(A+E)$ follows

$$-PA^{-1}E\hat{P} = P\hat{P} - PA^{-1}(A+E)\hat{P} = \sin\Theta - PA^{-1}(A+E)\hat{P}.$$

Again, using the fact that $PA = PAP$ and $(A+E)\hat{P} = \hat{P}(A+E)\hat{P}$ one obtains

$$\begin{aligned}-PA^{-1}E\hat{P} &= \sin\Theta - PA^{-1}\sin\Theta(A+E)\hat{P}\\ &= PA^{-1}PA\sin\Theta - PA^{-1}\sin\Theta(A+E)\hat{P}\\ &= PA^{-1}(PA\sin\Theta - \sin\Theta(A+E)\hat{P})\end{aligned}$$

and $\sin\Theta = P\sin\Theta\hat{P}$ implies

$$\|A^{-1}E\| \geqslant \|PA^{-1}E\hat{P}\| \geqslant \mathrm{relsep}\|\sin\Theta\|. \qquad \square$$

Thus, the subspace $\mathscr{S}$ is insensitive to relative perturbations $A^{-1}E$ if the relative separation is large. The derivation of the relative bound is very similar to the derivation of the absolute bound. In this sense, there is nothing special about a relative bound.

When the perturbed subspace has dimension one, the absolute bound implies the relative bound.

Theorem 3.3 (Ipsen [18, Theorem 3.3]). *If $\hat{\mathscr{S}}$ has dimension one then Theorem* 3.1 *implies Theorem* 3.2.

Proof. Since $\hat{\mathscr{S}}$ has dimension one, $\hat{X}$ consists of only one column, and $\hat{B}$ is a scalar. Hence one can write $(A+E)\hat{x} = \hat{\lambda}\hat{x}$. Using $\hat{P} = \hat{x}\hat{x}^*/\hat{x}^*\hat{x}$ and $PZ\hat{P} = Z$, Theorem 3.1 can be expressed as

$$\|\sin\Theta\| \leqslant \|E\|/\mathrm{abssep} \quad \text{where } \mathrm{abssep} = \min_{\|Z\|=1} \|P(A - \hat{\lambda}I)Z\|$$

and Theorem 3.2 as

$$\|\sin\Theta\| \leqslant \|A^{-1}E\|/\mathrm{relsep} \quad \text{where } \mathrm{relsep} = \min_{\|Z\|=1} \|PA^{-1}(A - \hat{\lambda}I)Z\|.$$

The idea is to write $(A+E)\hat{x} = \hat{\lambda}\hat{x}$ as $(\tilde{A} + \tilde{E})\hat{x} = \hat{x}$, where $\tilde{A} \equiv \hat{\lambda}A^{-1}$, and $\tilde{E} \equiv -A^{-1}E$. Note that $\tilde{A}$ and $\tilde{A} + \tilde{E}$ are associated with the same projectors P and $\hat{P}$, respectively, as A and $A+E$.

Theorem 3.1 implies Theorem 3.2 because applying the absolute bound to $(\tilde{A} + \tilde{E})\hat{x} = 1\cdot\hat{x}$ yields the relative bound. In particular, the norm in abssep is

$$\|P(\tilde{A} - 1\cdot I)Z\| = \|P(\hat{\lambda}A^{-1} - I)Z\| = \|PA^{-1}(A - \hat{\lambda}I)Z\|,$$

which is equal to the norm in relsep. $\square$

Since the relative bound is derived by means of the absolute bound one cannot necessarily conclude that relative perturbation bounds are stronger than absolute bounds. However, there are particular matrices and classes of perturbations, where relative bounds can be much sharper than absolute bounds.

Example 3.1 (*Ipsen* [18, *Example* 2]). Let $k>0$ and

$$A=\begin{pmatrix}10^{-k} & & \\ & 2\cdot 10^{-k} & \\ & & 10^{k}\end{pmatrix}.$$

Suppose $\mathscr{S}=\text{range}(1\ 0\ 0)^{\mathrm{T}}$ is approximated by the subspace associated with the smallest eigenvalue $\hat{\lambda}=10^{-k}$ of

$$A+E=\begin{pmatrix}10^{-k} & & \\ \varepsilon 10^{-k} & 2\cdot 10^{-k} & \\ \varepsilon 10^{k} & \varepsilon 10^{k} & 10^{k}\end{pmatrix},$$

where $\varepsilon>0$. In this case,

$$\mathscr{S}^{\perp}=\text{range}\begin{pmatrix}0 & 0\\ 1 & 0\\ 0 & 1\end{pmatrix},\qquad P=\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

The absolute bound contains

$$P(A-\hat{\lambda}I)=\begin{pmatrix}0 & 0 & 0\\ 0 & 10^{-k} & 0\\ 0 & 0 & 10^{k}-10^{-k}\end{pmatrix}.$$

Hence, in the two-norm abssep $\approx 10^{-k}$. Since $\|E\|_2\approx\varepsilon 10^{k}$, the absolute bound is

$$\|\sin\Theta\|_2\leqslant\|E\|_2/\text{abssep}\approx\varepsilon 10^{2k}.$$

In contrast, the relative bound contains

$$PA^{-1}(A-\hat{\lambda}I)=\begin{pmatrix}0 & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & 0 & 1-10^{-2k}\end{pmatrix}.$$

Hence, in the two-norm relsep ≈ 1. Since $\|A^{-1}E\|_2\approx\varepsilon$, the relative bound is

$$\|\sin\Theta\|_2\leqslant\|A^{-1}E\|_2/\text{relsep}\approx\varepsilon.$$

In this case the relative bound is sharper by a factor of 10^{2k} than the absolute bound. □

In general, it is not known, under which circumstances a relative bound is better than an absolute bound, and which type of relative bound is the tightest for a given matrix and perturbation.

So far, we have considered bounds between two subspaces that make no reference to any basis. From a computational point of view, however, this may not be useful. This is why from now on we express subspace bounds in terms of specified bases. Such bounds turn out to be weaker, as they are derived by bounding from below abssep in Theorem 3.1 and relsep in Theorem 3.2.

Let Y and $\hat{X}$ be respective bases for $\mathscr{S}^{\perp}$ and $\hat{\mathscr{S}}$, that is,

$$Y^{*}A=\Lambda Y^{*}\quad\text{where } \mathscr{S}^{\perp}=\text{range}(Y)$$

and

$$(A+E)\hat{X}=\hat{X}\hat{\Lambda}\quad\text{where } \hat{\mathscr{S}}=\text{range}(\hat{X})$$

for some matrices Λ and $\hat{\Lambda}$. This means, the eigenvalues of $\hat{\Lambda}$ are the eigenvalues associated with the perturbed subspace $\hat{\mathscr{S}}$, while the eigenvalues of Λ are the eigenvalues associated with the exact subspace in which we are *not* interested, because the associated left subspace is orthogonal to the desired subspace $\mathscr{S}$. However, the separation between the eigenvalues of Λ and $\hat{\Lambda}$ determines the quality of the perturbed subspace $\hat{\mathscr{S}}$. Denote by $\kappa(Y) \equiv \|Y\|_2 \|Y^\dagger\|_2$ the two-norm condition number with respect to inversion. The absolute bound in Theorem 3.1 can be weakened to [18, (4.2)]

$$\|\sin\Theta\| \leqslant \kappa(Y)\kappa(\hat{X})\|E\|/\mathrm{abssep}(\Lambda,\hat{\Lambda}), \tag{3.1}$$

where

$$\mathrm{abssep}(\Lambda,\hat{\Lambda}) \equiv \min_{\|Z\|=1} \|\Lambda Z - Z\hat{\Lambda}\|.$$

When A is nonsingular, the relative bound in Theorem 3.2 can be weakened to [18, (4.3)]

$$\|\sin\Theta\| \leqslant \kappa(Y)\kappa(\hat{X})\|A^{-1}E\|/\mathrm{relsep}(\Lambda,\hat{\Lambda}), \tag{3.2}$$

where

$$\mathrm{relsep}(\Lambda,\hat{\Lambda}) \equiv \min_{\|Z\|=1} \|\Lambda^{-1}(\Lambda Z - Z\hat{\Lambda})\|.$$

Unfortunately, bounds (3.1) and (3.2) contain a quantity in which we are not really interested, $\kappa(Y)$, the conditioning of a basis for $\mathscr{S}^\perp$. Usually, Y is not explicitly specified, and we have some freedom of choice here. There are two simple options. Either choose Y as a basis of Schur vectors (then Y has orthonormal columns and $\kappa(Y)=1$), or choose Y as a basis of Jordan vectors (then Λ is diagonal when A is diagonalizable). We make the later choice for diagonalizable matrices, so that absolute and relative separations can be expressed in terms of eigenvalues. For normal and Hermitian matrices, fortunately, the two choices coincide.

3.2. *Diagonalizable matrices*

Relative and absolute bounds for eigenspaces of diagonalizable matrices are expressed in terms of eigenvalues and conditioning of eigenvector bases.

Let $\mathscr{S}$ and $\hat{\mathscr{S}}$ be respective eigenspaces for diagonalizable matrices A and $A+E$, and let the columns of Y and $\hat{X}$ be respective bases for $\mathscr{S}^\perp$ and $\hat{\mathscr{S}}$. That is

$$\mathscr{S}^\perp = \mathrm{range}(Y), \quad \hat{\mathscr{S}} = \mathrm{range}(\hat{X})$$

and

$$Y^*A = \Lambda Y^*, \quad (A+E)\hat{X} = \hat{X}\hat{\Lambda},$$

where Λ and $\hat{\Lambda}$ are diagonal. We (ab)use the notation

$$\min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda-\hat{\lambda}| \quad \text{and} \quad \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}$$

to mean that the minima range over all diagonal elements λ of Λ and all diagonal elements $\hat{\lambda}$ of $\hat{\Lambda}$.

Theorem 3.4. *If A and $A+E$ are diagonalizable then*

$$\|\sin\Theta\|_F \leqslant \kappa(Y)\kappa(\hat{X})\ \|E\|_F \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda-\hat{\lambda}|.$$

If, in addition, A is nonsingular, then

$$\|\sin\Theta\|_F \leqslant \kappa(Y)\kappa(\hat{X})\ \|A^{-1}E\|_F \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}.$$

Proof. The absolute bound follows from (3.1) and the fact that $\mathrm{abssep}_F(\Lambda,\hat{\Lambda}) = \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda-\hat{\lambda}|$ [34, p. 245, Problem 3]. Regarding the relative bound, the norm in $\mathrm{relsep}(\Lambda,\hat{\Lambda})$ can be bounded by

$$\|Z-\Lambda^{-1}Z\hat{\Lambda}\|_F^2 = \sum_{i,j} \left|1-\frac{\hat{\lambda}_j}{\lambda_i}\right|^2 |z_{ij}|^2 \geqslant \min_{i,j} \left|1-\frac{\hat{\lambda}_j}{\lambda_i}\right|^2 \|Z\|_F^2.$$

Now use $\mathrm{relsep}_F(\Lambda,\hat{\Lambda}) \geqslant \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}$ in (3.2). □

Thus, the eigenspace $\mathscr{S}$ is insensitive to absolute (relative) perturbations if the eigenvector bases are well-conditioned and if the perturbed eigenvalues are well-separated in the absolute (relative) sense from the undesired exact eigenvalues.

In the particular case when $\hat{\mathscr{S}}$ has dimension 1, the absolute bound in Theorem 3.4 reduces to [13, Theorem 3.1], see also Theorem 4.1.

Bounds similar to the Frobenius norm bounds in Theorem 3.4 can be derived for the two-norm. This is done either by bounding the Frobenius norm in terms of the two-norm and inheriting a factor of $\sqrt{n}$ in the bound, where n is the order of A [18, Corollary 5.2], or by assuming that all eigenvalues of one matrix (Λ or $\hat{\Lambda}$) are smaller in magnitude than all eigenvalues of the other matrix [18, Theorem 5.3].

When A and $A+E$ are normal, the condition numbers for the eigenvector bases equal one, and the Frobenius norm bounds in Theorem 3.4 simplify.

Corollary 3.5. *If A and $A+E$ are normal then*

$$\|\sin\Theta\|_F \leqslant \|E\|_F \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda-\hat{\lambda}|.$$

If, in addition, A is non-singular, then

$$\|\sin\Theta\|_F \leqslant \|A^{-1}E\|_F \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}.$$

Now the sensitivity of the subspace to absolute (relative) perturbations depends solely on the absolute (relative) eigenvalue separation. The absolute bound represents one of Davis and Kahan's $\sin\Theta$ Theorems [4, Section 6]; [5, Section 2].

In particular, the above bounds hold for Hermitian matrices. However, the relative perturbation $A^{-1}E$ is, in general, not Hermitian. By expressing the relative perturbation differently, one can obtain Hermitian perturbations. This is done in the following sections, where things become more complex because we demand structure from relative perturbations. For instance, when relative perturbations are restricted to be Hermitian, subspaces of indefinite Hermitian matrices appear to be more sensitive than those of definite matrices.

3.3. Hermitian positive-definite matrices

Relative bounds with Hermitian perturbations are derived for eigenspaces of Hermitian positive-definite matrices. We start by discussing positive-definite matrices because it is easy to construct relative perturbations that are Hermitian. Construction of Hermitian relative perturbations for indefinite matrices is more intricate, but the derivations are often guided by those for definite matrices.

In contrast to the preceding results, one would like to express relative perturbations for Hermitian matrices as $A^{-1/2}EA^{-1/2}$, where $A^{1/2}$ is a square-root of A. The nice thing about Hermitian positive-definite matrices A is that one can choose $A^{1/2}$ to be Hermitian. Hence $A^{-1/2}EA^{-1/2}$ remains Hermitian whenever E is Hermitian.

Let $\mathscr{S}$ and $\hat{\mathscr{S}}$ be respective eigenspaces for Hermitian positive-definite matrices A and $A+E$, and let the columns of Y and $\hat{X}$ be respective orthonormal bases for $\mathscr{S}^{\perp}$ and $\hat{\mathscr{S}}$. That is

$$\mathscr{S}^{\perp} = \mathrm{range}(Y), \quad \hat{\mathscr{S}} = \mathrm{range}(\hat{X})$$

and

$$Y^*A = \Lambda Y^*, \quad (A+E)\hat{X} = \hat{X}\hat{\Lambda},$$

where Λ and $\hat{\Lambda}$ are diagonal with positive diagonal elements. Since Y and $\hat{X}$ have orthonormal columns, $\|\sin\Theta\| = \|Y^*\hat{X}\|$.

The derivation of the relative bound below was inspired by the proof of [26, Theorem 1].

Theorem 3.6 (Londrè and Rhee [22, Theorem 1], Li [21, Theorem 3.3]). *If A and $A+E$ are Hermitian positive-definite, and if $\eta_2 \equiv \|A^{-1/2}EA^{-1/2}\|_2 < 1$ then*

$$\|\sin\Theta\|_{\mathrm{F}} \leqslant \frac{\eta_{\mathrm{F}}}{\sqrt{1-\eta_2}} \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{\lambda\hat{\lambda}}},$$

where $\eta_{\mathrm{F}} \equiv \|A^{-1/2}EA^{-1/2}\|_{\mathrm{F}}$.

Proof. Multiply $(A+E)\hat{X} = \hat{X}\hat{\Lambda}$ on the left by Y^* and set $S \equiv Y^*\hat{X}$,

$$\Lambda S - S\hat{\Lambda} = -Y^*E\hat{X} = -\Lambda^{1/2}\, W\, \hat{\Lambda}^{1/2},$$

where $W \equiv -\Lambda^{-1/2}\, Y^*E\hat{X}\, \hat{\Lambda}^{-1/2}$. Element (i,j) of the equation is

$$s_{ij} = -W_{ij} \Big/ \frac{\lambda_i - \hat{\lambda}_j}{\sqrt{\lambda_i\hat{\lambda}_j}},$$

where λ_i and $\hat{\lambda}_j$ are respective diagonal elements of Λ and $\hat{\Lambda}$. Summing up all elements gives

$$\|\sin\Theta\|_{\mathrm{F}} = \|S\|_{\mathrm{F}} \leqslant \|W\|_{\mathrm{F}} \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{\lambda\hat{\lambda}}}.$$

From

$$W = Y^* A^{-1/2} E (A+E)^{-1/2} \hat{X} = Y^* \, A^{-1/2} E A^{-1/2} A^{1/2} (A+E)^{-1/2} \hat{X}$$

follows

$$\|W\|_F \leqslant \eta_F \, \|A^{1/2}(A+E)^{-1/2}\|_2.$$

Positive-definiteness is crucial for bounding $\|A^{1/2}(A+E)^{-1/2}\|_2$. Since $A^{1/2}$ and $(A+E)^{1/2}$ are Hermitian,

$$\|A^{1/2}(A+E)^{-1/2}\|_2^2 = \|A^{1/2}(A+E)^{-1}A^{1/2}\|_2 = \|(I + A^{-1/2}EA^{1/2})^{-1}\|_2$$

$$\leqslant \frac{1}{1-\eta_2}. \qquad \square$$

Thus, the eigenspace $\mathscr{S}$ is insensitive to relative perturbations $A^{-1/2}EA^{-1/2}$ if the relative separation between perturbed eigenvalues and the undesirable exact eigenvalues is large. Since the relative perturbation $A^{-1/2}EA^{-1/2}$ in Theorem 3.6 is different from the preceding perturbation $A^{-1}E$, so is the relative eigenvalue separation. However, this is of little consequence: If one measure of relative eigenvalue separation is small, so are all others [20, Section 2]; [25, Section 1]. The above bound holds more generally for unitarily invariant norms [21, Theorem 3.4].

With regard to related developments, a bound on $|A^{1/2}(A+E)^{-1/2} - I\|_2$ is derived in [24]. Relative eigenvector bounds for the hyperbolic eigenvalue problem $Ax = \lambda Jx$, where A is Hermitian positive-definite and J is a diagonal matrix with diagonal entries of magnitude one are given in [31, Section 3.2], with auxiliary results in [33].

A relative perturbation of the form $A^{-1/2}EA^{-1/2}$ not only has the advantage that it is Hermitian, it is also invariant under grading, when both A and E are graded in the same way. This is discussed in the next section.

3.4. Graded Hermitian positive-definite matrices

It is shown that the relative perturbations $A^{-1/2}EA^{-1/2}$ from the previous section are invariant under grading. By 'grading' (or 'scaling') [1, Section 2]; [25, Section 1] we mean the following: There exists a nonsingular matrix D such that $A = D^*MD$ where M is in some sense 'better-behaved' than A.

Lemma 3.7 (Eisenstat and Ipsen [14, Corollary 3.4], Mathias [25, Lemma 2.2]). *If $A = D^*MD$ is positive definite and $E = D^*FD$ then*

$$\|A^{-1/2}EA^{-1/2}\| = \|M^{-1/2}FM^{-1/2}\|.$$

Proof. We reproduce here the proof of [19, Corollary 2.13]. Because A is Hermitian positive-definite, it has a Hermitian square-root $A^{1/2}$. Hence $A^{-1/2}EA^{-1/2}$ is Hermitian, and the norm is an eigenvalue,

$$\|A^{-1/2}EA^{-1/2}\| = \max_{1 \leqslant j \leqslant n} |\lambda_j(A^{-1/2}EA^{-1/2})|.$$

Now comes the trick. Since eigenvalues are preserved under similarity transformations, we can reorder the matrices in a circular fashion until all grading matrices have cancelled each other out,

$$\lambda_j(A^{-1/2}EA^{-1/2}) = \lambda_j(A^{-1}E) = \lambda_j(D^{-1}\, M^{-1}F\, D) = \lambda_j(M^{-1}F)$$
$$= \lambda_j(M^{-1/2}FM^{-1/2}).$$

At last recover the norm,

$$\max_{1\leqslant j\leqslant n} |\lambda_j(M^{-1/2}FM^{-1/2})| = \|M^{-1/2}FM^{-1/2}\|. \qquad \square$$

Application of Lemma 3.7 to Theorem 3.6 demonstrates that the relative perturbations do not depend on the grading matrix D.

Corollary 3.8 (Li [21, Theorem 3.3]). *If $A = D^*MD$ and $A + E = D^*(M + F)D$ are Hermitian positive-definite, where D is nonsingular, and if*

$$\eta_2 \equiv \|M^{-1/2}FM^{-1/2}\|_2 < 1$$

then

$$\|\sin\Theta\|_{\mathrm{F}} \leqslant \frac{\eta_{\mathrm{F}}}{\sqrt{1-\eta_2}} \Bigg/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{\lambda\hat{\lambda}}},$$

where $\eta_{\mathrm{F}} \equiv \|M^{-1/2}FM^{-1/2}\|_{\mathrm{F}}$.

Again, the above bound holds more generally for unitarily invariant norms [21, Theorem 3.4].

There are other relative bounds for Hermitian positive-definite matrices D^*MD that exploit grading in the error D^*FD.

- Component-wise first-order bounds on the difference between perturbed and exact eigenvectors, containing the perturbation $\|M^{-1/2}FM^{-1/2}\|_2$, a relative gap, as well as eigenvalues and diagonal elements of M [25, Section 3].
- Component-wise exact bounds with the same features as above [22, Theorem 4].
- Norm-wise and component-wise first-order bounds on the difference between exact and perturbed eigenvectors, containing eigenvalues of M and $\|F\|_2$ [8, Section 2]. Here D is diagonal so that all diagonal elements of M are equal to one.

The next section shows how to deal with indefinite matrices, first without and then with grading.

3.5. Hermitian indefinite matrices

The bound for positive-definite Hermitian matrices in Section 3.3 is extended to indefinite matrices, however with a penalty. The penalty comes about, it appears, because the relative perturbation is asked to be Hermitian.

To understand the penalty, it is necessary to introduce polar factors and J-unitary matrices. Let A be Hermitian matrix with eigendecomposition $A = V\Omega V^*$ and denote by $|\Omega|$ the diagonal matrix whose diagonal elements are the absolute values of the diagonal elements in Ω. The generalization of this absolute value to non-diagonal matrices is the Hermitian positive-definite polar factor (or spectral absolute value [37, Section 1]) of A, $|A| \equiv V\,|\Omega|\,V^*$. When A happens to be positive-definite then $|A| = A$. Note that the polar factor $|A|$ has the same eigenvectors as A.

The J in the J-unitary matrices comes from the inertia of A. Write an eigendecomposition of A

$$A = V\Omega V^* = V\,|\Omega|^{1/2} J |\Omega|^{1/2} V^*,$$

where J is a diagonal matrix with ± 1 on the diagonal that reflects the inertia of A. A matrix Z with $ZJZ^* = J$ is called J-unitary. When A is definite then $J = \pm I$ is a multiple of the identity, hence J-unitary matrices are just plain unitary. One needs J-unitary matrices to transform one decomposition of an indefinite matrix into another. For instance, suppose one has two decompositions $A = Z_1 J Z_1^* = Z_2 J Z_2^*$. Then there exists a J-unitary matrix Z that transforms Z_1 into Z_2. That is,

$$Z_1 = Z_2 Z, \quad ZJZ^* = J.$$

One such matrix is simply $Z = Z_2^{-1} Z_1$.

Now we are ready to extend Theorem 3.6.

Theorem 3.9 (Simpler Version of Theorem 2 in Truhar and Slapničar, [36]). *If A and $A+E$ are Hermitian with the same inertia, and if $\eta_2 \equiv \|\,|A|^{-1/2} E |A|^{-1/2}\|_2 < 1$ then*

$$\|\sin\Theta\|_F \leqslant \|Z\|_2 \frac{\eta_F}{\sqrt{1-\eta_2}} \bigg/ \min_{\lambda\in\Lambda, \hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{\sqrt{\lambda\hat{\lambda}}},$$

where $\eta_F \equiv \|\,|A|^{-1/2} E |A|^{-1/2}\|_F$ and Z is J-unitary and defined in the proof below.

Proof. The proof is very similar to that of Theorem 3.6. Multiply $(A+E)\hat{X} = \hat{X}\hat{\Lambda}$ on the left by Y^* and set $S \equiv Y^*\hat{X}$,

$$\Lambda S - S\hat{\Lambda} = -Y^* E \hat{X} = -|\Lambda|^{1/2} W |\hat{\Lambda}|^{1/2},$$

where $W \equiv -|\Lambda|^{-1/2} Y^* E \hat{X} |\hat{\Lambda}|^{-1/2}$. Element (i,j) of the equation is

$$s_{ij} = -W_{ij} \bigg/ \frac{\lambda_i - \hat{\lambda}_j}{\sqrt{|\lambda_i \hat{\lambda}_j|}},$$

where λ_i and $\hat{\lambda}_j$ are respective diagonal elements of Λ and $\hat{\Lambda}$. Summing up all elements gives

$$\|\sin\Theta\|_F = \|S\|_F \leqslant \|W\|_F \bigg/ \min_{\lambda\in\Lambda, \hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{\sqrt{|\lambda\hat{\lambda}|}}.$$

From

$$W = Y^* |A|^{-1/2} E |A+E|^{-1/2} \hat{X} = Y^* |A|^{-1/2} E |A|^{-1/2} |A|^{1/2} |A+E|^{-1/2} \hat{X}$$

follows

$$\|W\|_F \leqslant \eta_F \|\,|A|^{1/2} |A+E|^{-1/2}\|_2.$$

Bounding $\|\,|A|^{1/2}|A+E|^{-1/2}\|_2$ requires more work than in the positive-definite case. The eigendecompositions $A = V\Omega V^*$ and $A+E = \hat{V}\hat{\Omega}\hat{V}^*$ lead to two decompositions for $A+E$,

$$A + E = \hat{V} |\hat{\Omega}|^{1/2} J |\hat{\Omega}|^{1/2} \hat{V}^*$$

and

$$\begin{aligned}A + E &= V|\Omega|^{1/2}J|\Omega|^{1/2}V^* + E \\ &= V|\Omega|^{1/2}(J + |\Omega|^{-1/2}V^*EV|\,|\Omega|^{-1/2})|\Omega|^{1/2}V^* \\ &= (V|\Omega|^{1/2}Q|\Delta|^{1/2})J(|\Delta|^{1/2}Q^*|\Omega|^{1/2}V^*),\end{aligned}$$

where

$$J + |\Omega|^{-1/2}\ V^*EV|\,|\Omega|^{-1/2} = Q\Delta Q^*$$

is an eigendecomposition with the same inertia as $A + E$ since we got there via a congruence transformation. To summarize the two expressions

$$A + E = Z_1JZ_1^* = Z_2JZ_2^*,$$

where

$$Z_1 \equiv \hat{V}|\hat{\Omega}|^{1/2}, \quad Z_2 \equiv V|\Omega|^{1/2}Q|\Delta|^{1/2}.$$

As explained above there exists a J-unitary matrix Z such that $Z_2 = Z_1Z$. Use this in

$$\|\,|A|^{1/2}|A+E|^{-1/2}\|_2 = \|\,|\Omega|^{1/2}V^*\hat{V}|\hat{\Omega}|^{-1/2}\|_2 = \|\,|\Omega|^{1/2}V^*\hat{V}Z_1^{-*}\|$$

to obtain

$$\|\,|A|^{1/2}|A+E|^{-1/2}\|_2 = \|\,|\Delta|^{-1/2}Z^*\|_2 \leqslant \|\Delta^{-1}\|^{1/2}\|Z\|_2$$

since Δ is a diagonal matrix. It remains to bound $\|\Delta^{-1}\|_2$,

$$\begin{aligned}\|\Delta^{-1}\|_2 &= \|(J + |\Omega|^{-1/2}V^*EV|\,|\Omega|^{-1/2})^{-1}\| \\ &= \|(I + J|\Omega|^{-1/2}\ V^*EV|\,|\Omega|^{-1/2})^{-1}\| \leqslant \frac{1}{1-\eta_2}. \qquad \square\end{aligned}$$

The bound in Theorem 3.9 looks similar to the bound in Theorem 3.6. But the square-roots in η_2 and η_F now contain polar factors, and the relative eigenvalue separation has absolute values under the square-root. Moreover, there is an additional factor $\|Z\|$, that's the penalty. In the lucky case when A happens to be positive-definite, Z is unitary and Theorem 3.9 reduces to Theorem 3.6. When A is indefinite, the eigenspace sensitivity can be magnified by the norm of the J-unitary matrix, which in some sense reflects the deviation of A from definiteness.

At this point it is not known how large $\|Z\|$ can be, under which circumstances it will be large or small, and how much it really contributes to the sensitivity of a subspace. A quantity corresponding to $\|Z\|$ in [36] is bounded in terms of $\|A^{-1}\|$ and a graded polar factor of A. Preliminary experiments in [36, Sections 4 and 5] suggest that $\|Z\|$ does not grow unduly. At present, we do not yet have a good understanding of why a subspace of an indefinite Hermitian matrix should be more sensitive to Hermitian relative perturbations than a subspace of a definite matrix.

Not all relative bounds for Hermitian matrices necessarily look like the one above. For instance, there are relative bounds specifically geared towards real symmetric tridiagonal matrices. The cosine between two Ritz vectors associated with an eigenvalue cluster of a real, symmetric tridiagonal matrix can be expressed in terms of a relative gap [27, Section 5]. Perturbations of the LDL^{T} decomposition of a real, symmetric tridiagonal matrix lead to relative bounds on the tangent between eigenvectors, and an eigenvector condition number that depends on all eigenvalues, not just a single eigenvalue separation [28, Section 10].

Like a tridiagonal matrix, one can decompose any Hermitian matrix as $A = G^*JG$, where J is a diagonal matrix with diagonal entries ± 1. The norm-wise perturbation of a spectral projector induced by perturbations of the factor G can be bounded in terms of a relative eigenvalue separation [35, (12)]; [32, Theorem 1].

3.6. Graded indefinite Hermitian matrices

The bound for graded positive-definite matrices from Section 3.4 is extended to graded indefinite matrices,

Fortunately, this requires only a slight modification in the proof of Theorem 3.9.

Theorem 3.10 (Simpler Version of Theorem 2 in Truhar and Slapničar [36]). *If $A=D^*MD$ and $A+E = D^*(M+F)D$ are Hermitian, where D is nonsingular, and if*

$$\eta_2 \equiv \| \, |M|^{-1/2} F |M|^{-1/2} | \, \|_2 < 1$$

then

$$\|\sin \Theta\|_F \leqslant \|Z\| \frac{\eta_F}{\sqrt{1-\eta_2}} \Big/ \min_{\lambda\in\Lambda, \hat\lambda\in\hat\Lambda} \frac{|\lambda-\hat\lambda|}{\sqrt{\lambda\hat\lambda}},$$

where $\eta_F \equiv \| \, |M|^{-1/2} F |M|^{-1/2} | \, \|_F$, and Z is J-unitary.

Proof. As in the proof of Theorem 3.9 derive

$$\|\sin \Theta\|_F = \|S\|_F \leqslant \|W\|_F \Big/ \min_{\lambda\in\Lambda, \hat\lambda\in\hat\Lambda} \frac{|\lambda-\hat\lambda|}{\sqrt{|\lambda\hat\lambda|}}.$$

To bound $\|W\|_F$, represent A and $A+E$ in terms of D and eigendecompositions of M and $M+F$, respectively. The scaling matrices D then cancel out with the scaling matrices in the error $E = D^*FD$. □

A quantity corresponding to $\|Z\|$ in [36] is bounded in terms of $\|A^{-1}\|$ and $\|D^*|M|D\|$.

Other relative bounds for indefinite Hermitian matrices that exploit grading include the following.

- Norm-wise and component-wise first-order bounds on the difference between exact and perturbed eigenvectors of real, symmetric scaled diagonally dominant matrices [1, Section 7].
- Bounds on the norm-wise difference between corresponding eigenvectors of Hermitian matrices $A = D^*MD$ and $D^*(M+F)D$ in terms of a relative gap, $\|F\|_2$ and an eigenvalue of a principal submatrix of M [17, Theorem 7]. This is an improvement over the bounds for symmetric scaled diagonally dominant matrices in [1, Section 7] and for positive-definite matrices in [8].

- Bounds on the cosines of angles between exact and perturbed eigenvectors of possibly singular Hermitian matrices [2, Section 4]. They can be applied to analyze the accuracy of subspaces in ULV down-dating [3].
- Bounds on the norm-wise perturbations in spectral projectors [29, Section 2], [35, (6), (7)], [37, Theorem 2.48]; [30].

3.7. Remarks

The sensitivity of invariant subspaces to absolute perturbations E and to relative perturbations $A^{-1}E$ is influenced by the same factors: conditioning of subspace bases, and separation of matrices associated with eigenvalues. When the matrices involved are Hermitian the sensitivity to absolute perturbations E is amplified by an absolute eigenvalue separation, and the sensitivity to relative perturbations $A^{-1}E$ by a relative eigenvalue separation. None of these two perturbations seems to care about whether the Hermitian matrices are definite or indefinite.

This changes when one restricts relative perturbations to be Hermitian as well, i.e., of the form $|A|^{-1/2}E|A|^{-1/2}$. Then subspaces of indefinite matrices appear to be more sensitive to these perturbations than those of definite matrices. This phenomenon is not yet completely understood. In particular, it is not clear how much the sensitivity can worsen for an indefinite matrix, and in what way the sensitivity depends on the indefiniteness of the matrix. In general, one does not completely understand how exactly the fine-structure of a matrix and a perturbation affect the sensitivity of subspaces.

There is another observation that has not been fully exploited yet either. Invariant subspaces do not change under shifts, i.e., A and the shifted matrix $A - \mu I$ have the same invariant subspaces. The condition numbers for the absolute perturbations are invariant under a shift, while those for relative perturbations are not. The question is, are there *optimal* shifts for computing subspaces, and what would 'optimal' mean in this context? In particular, one could shift a Hermitian matrix so it becomes positive-definite. Then the subspaces of the shifted matrix would look less sensitive to Hermitian relative perturbations. This approach is pursued to assess the sensivity of eigenvectors of factored real symmetric tridiagonal matrices to relative perturbations in the factors in [9, Section 5], [28, Section 10], and used to compute the eigenvectors in [10]. The approach based on shifting a matrix before evaluating sensitivity and computing subspaces deserves more investigation for general, Hermitian matrices.

Now we consider a different type of perturbation.

4. Multiplicative perturbations

The perturbed subspace $\mathscr{S}$ is interpreted as an exact subspace of a perturbed matrix D_1AD_2, where D_1 and D_2 are nonsingular. Relative and absolute bounds on $\|\sin\Theta\|$ are presented for diagonalizable and Hermitian matrices.

When $D_2 = D_1^{-1}$, the perturbed matrix D_1AD_2 is just a similarity transformation of A, which means that A and D_1AD_2 have the same eigenvalues. When $D_2 = D_1^*$ then D_1AD_2 is a congruence transformation of A, which means that A and D_1AD_2 have the same inertia when A is Hermitian.

Since the nonsingularity of D_1 and D_2 forces A and D_1AD_2 to have the same rank, multiplicative perturbations are more restrictive than additive perturbations.

Multiplicative perturbations can be used, for instance, to represent component-wise perturbations of real bidiagonal matrices and of real symmetric tridiagonal matrices with zero diagonal [1, p. 770], [12, Section 4], [19, Example 5.1]. This is exploited in [28, Section 4], where the relative sensitivity of eigenvalues and eigenvectors of real symmetric tridiagonal matrices with regard to perturbations in the factors of a LDL^{T} factorization is analyzed. Since L is bidiagonal, a component-wise perturbation of L can be represented as D_1LD_2.

In a different application illustrated below, multiplicative perturbations represent deflation in block triangular matrices.

Example 4.1 (Eisenstat and Ipsen [12, *Theorem* 5.2])**.** The off-diagonal block in the block triangular matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ & A_{22} \end{pmatrix}$$

is to be eliminated, making the deflated matrix

$$\begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix}$$

block diagonal. When A_{11} is nonsingular one can factor

$$\begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ & A_{22} \end{pmatrix} \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ & I \end{pmatrix}.$$

Therefore, the deflated matrix represents a multiplicative perturbation D_1AD_2, where $D_1 = I$ and

$$D_2 = \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ & I \end{pmatrix}.$$

Similarly, when A_{22} is nonsingular one can factor

$$\begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix} = \begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ & A_{22}. \end{pmatrix}.$$

In this case the deflated matrix represents a multiplicative perturbation D_1AD_2, where $D_2 = I$ and

$$D_2 = \begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ & I \end{pmatrix}. \qquad \square$$

4.1. Diagonalizable matrices

A bound is presented between a perturbed one-dimensional eigenspace and an eigenspace of a diagonalizable matrix.

Suppose A is diagonalizable and $\hat{\mathscr{S}} = \mathrm{range}(\hat{x})$, where

$$D_1AD_2\hat{x} = \hat{\lambda}\hat{x}, \quad \|\hat{x}\|_2 = 1$$

for some nonsingular D_1 and D_2, where D_1AD_2 is not necessarily diagonalizable. In this section we explicitly choose $\mathscr{S}$ to be the eigenspace associated with all eigenvalues of A closest to $\hat{\lambda}$. The remaining, further away eigenvalues form the diagonal elements of the diagonal matrix Λ, i.e.,

$$\min_{\lambda\in\Lambda} |\lambda - \hat{\lambda}| > \min_i |\lambda_i(A) - \hat{\lambda}|.$$

Then $\mathscr{S}^\perp$ is the left invariant subspace of A associated with the eigenvalues in Λ. Let the columns of Y be a basis for $\mathscr{S}^\perp$, so $\mathscr{S}^\perp = \mathrm{range}(Y)$ and

$$Y^*A = \Lambda Y^*.$$

In the theorem below the residual of $\hat{x}$ and $\hat{\lambda}$ is

$$r \equiv (A - \hat{\lambda}I)\hat{x}.$$

Theorem 4.1 (Eisenstat and Ipsen [13, Theorem 4.3]). *If A is diagonalizable then*

$$\|\sin\Theta\|_2 \leqslant \kappa(Y)\|r\|_2 / \min_{\lambda\in\Lambda} |\lambda - \hat{\lambda}|.$$

If, in addition, D_1 and D_2 are nonsingular then

$$\|\sin\Theta\|_2 \leqslant \kappa(Y)\min\{\alpha_1,\alpha_2\} / \min_{\lambda\in\Lambda} \frac{|\lambda - \hat{\lambda}|}{|\hat{\lambda}|} + \|I - D_2\|_2,$$

where

$$\alpha_1 \equiv \|D_1^{-1} - D_2\|_2, \quad \alpha_2 \equiv \|I - D_1^{-1}D_2^{-1}\|_2.$$

Proof. To derive the absolute bound, multiply $r = (A - \hat{\lambda}I)\hat{x}$ by Y^* and use $Y^*A = Y^*\Lambda$,

$$Y^*\hat{x} = (\Lambda - \hat{\lambda}I)^{-1}Y^*r.$$

With P being the orthogonal projector onto $\mathscr{S}^\perp = \mathrm{range}(Y)$ one gets

$$P\hat{x} = (Y^\dagger)^*Y^*\hat{x} = (Y^\dagger)^*(\Lambda - \hat{\lambda}I)^{-1}Y^*r.$$

From $\|\hat{x}\|_2 = 1$ follows

$$\|\sin\Theta\|_2 = \|P\hat{x}\|_2 \leqslant \kappa(Y)\|(\Lambda - \hat{\lambda}I)^{-1}\|_2\|r\|_2.$$

To derive the relative bound, we will use the absolute bound. Multiply $(D_1AD_2)\hat{x} = \hat{\lambda}\hat{x}$ by D_1^{-1} and set $z \equiv D_2\hat{x}/\|D_2\hat{x}\|$,

$$Az = \hat{\lambda}D_1^{-1}D_2^{-1}z.$$

The residual for $\hat{\lambda}$ and z is

$$f \equiv Az - \hat{\lambda}z = \hat{\lambda}(D_1^{-1}D_2^{-1} - I)z = \hat{\lambda}(D_1^{-1} - D_2)\hat{x}/\|D_2\hat{x}\|_2.$$

Hence

$$\|f\|_2 \leqslant |\hat{\lambda}|\alpha_2, \quad \|f\|_2 \leqslant |\hat{\lambda}|\alpha_1/\|D_2\hat{x}\|_2.$$

The idea is to first apply the absolute bound to the residual f and then make an adjustment from z to $\hat{x}$. Since f contains $\hat{\lambda}$ as a factor we will end up with a relative bound.

Applying the absolute bound to f gives

$$\|\sin \Phi\|_2 \leqslant \kappa(Y)\|f\|_2 / \min_{\lambda\in\Lambda} |\lambda - \hat{\lambda}|,$$

where Φ represents the angle between z and $\mathscr{S}$. To make the adjustment from z to $\hat{x}$ use the fact that [13, Lemma 4.2]

$$\|\sin \Theta\|_2 \leqslant \|\sin \Phi\|_2 + \|D_2 - I\|_2$$

and

$$\|\sin \Theta\|_2 \leqslant \|D_2\hat{x}\|_2 \|\sin \Phi\|_2 + \|D_2 - I\|_2.$$

Now put the first bound for $\|f\|_2$ into the first bound for $\|\sin \Theta\|_2$ and the second bound for $\|f\|_2$ into the second bound for $\|\sin \Theta\|_2$. □

The relative bound consists of two summands. The first summand represents the (absolute or relative) deviation of D_1 and D_2 from a similarity transformation, amplified by the eigenvector conditioning $\kappa(Y)$ and by the relative eigenvalue separation; while the second summand represents the (absolute and relative) deviation of the similarity transformation from the identity. The factor α_1 is an absolute deviation from similarity, while α_2 constitutes a relative deviation as

$$I - D_1^{-1}D_2^{-1} = (D_2 - D_1^{-1})D_2^{-1}$$

is a difference relative to D_2. Thus, for an eigenspace to be insensitive to multiplicative perturbations, the multiplicative perturbations must constitute a similarity transformation close to the identity.

Here again, as in Theorem 3.3, the relative bound is implied by the absolute bound. Also, when $\hat{\mathscr{S}}$ has dimension 1, the absolute bound in Theorem 4.1 implies the absolute bound in Theorem 3.4.

Example 4.2. Let us apply Theorem 4.1 to Example 4.1. Suppose $\hat{x}$ is a unit-norm eigenvector associated with an eigenvalue $\hat{\lambda}$ of the deflated, block-diagonal matrix.

First consider the case when A_{11} is nonsingular. Then $D_1 = I$ and

$$\alpha_1 = \alpha_2 = \|I - D_2\|_2 = \|A_{11}^{-1}A_{12}\|_2.$$

Hence

$$\|\sin \Theta\|_2 \leqslant \|A_{11}^{-1}A_{12}\|_2 \left(1 + 1/\min_{\lambda\in\Lambda} \frac{|\lambda - \hat{\lambda}|}{|\hat{\lambda}|}\right).$$

This means $\hat{x}$ is close to an eigenvector of A if $\|A_{11}^{-1}A_{12}\|_2$ is small compared to 1 and the relative eigenvalue separation. Hence, the matrix can be safely deflated without harming the eigenvector, if the leading diagonal block is 'large enough compared to' the off-diagonal block, and the meaning of 'large enough' is determined by the eigenvalue separation.

In the second case when A_{22} is nonsingular, one has $D_2 = I$. Hence $\|D_2 - I\|_2 = 0$, $\alpha_1 = \alpha_2 = \|A_{12}A_{22}^{-1}\|_2$, and

$$\|\sin \Theta\|_2 \leqslant \|A_{11}^{-1}A_{12}\|_2 / \min_{\lambda\in\Lambda} \frac{|\lambda - \hat{\lambda}|}{|\hat{\lambda}|}.$$

Now the matrix can be safely deflated without harming the eigenvector, if the trailing diagonal block is 'large enough compared to' the off-diagonal block. □

In some cases the first summand in the bound of Theorem 4.1 can be omitted.

Corollary 4.2 (Eisenstat and Ipsen [13, Corollary 4.4]). *If $D_1 = D_2^{-1}$ or $\hat{\lambda} = 0$, then*

$$\|\sin \Theta\|_2 \leqslant \|I - D_2\|_2.$$

Proof. First suppose $D_1 = D_2^{-1}$. Then $D_1 A D_2 \hat{x} = \hat{\lambda}\hat{x}$ implies $A D_2 \hat{x} = \hat{\lambda} D_2 \hat{x}$, i.e., $\hat{\lambda}$ and $D_2\hat{x}$ are an exact eigenpair of A. Since $\mathscr{S}$ is the eigenspace associated with all eigenvalues closest to $\hat{\lambda}$, we must have $D_2\hat{x} \in \mathscr{S}$. Hence $PD_2\hat{x} = 0$, where P is the orthogonal projector onto $\mathscr{S}^{\perp}$, and

$$\|\sin \Theta\|_2 = \|P\hat{x}\|_2 = \|P(D_2\hat{x} - \hat{x})\|_2 \leqslant \|I - D_2\|_2.$$

Now suppose $\hat{\lambda} = 0$. Then $D_1 A D_2 \hat{x} = 0 \cdot \hat{x}$ implies $D_2^{-1} A D_2 \hat{x} = 0 \cdot \hat{x}$, since D_1 and D_2 are nonsingular. Hence $\hat{\lambda}$ and $\hat{x}$ are an exact eigenpair of a similarity transformation of A, and we are back to the first case. □

In the case of similarity transformations $D_1 = D_2^{-1}$, the eigenspace angle is bounded by the relative deviation of D_2 from identity, without any amplification by $\kappa(Y)$ or by a relative gap. As a consequence, eigenvectors of diagonalizable matrices are well-conditioned when the perturbation is a similarity transformation. Similarly, in the case $\hat{\lambda} = 0$ it follows that null vectors of diagonalizable matrices are well-conditioned under multiplicative perturbations.

A different approach is sketched in [21, Remark 3.3] for deriving eigenspace bounds of diagonalizable matrices when both eigenspaces have the same dimension $\geqslant 1$.

4.2. Hermitian matrices

Two-norm and Frobenius norm bounds are presented for multiplicative perturbations that are congruence transformations.

When applied to Hermitian matrices, Theorem 4.1 simplifies. Remember that in this context the perturbed eigenspace has dimension one, $\hat{\mathscr{S}} = \mathrm{range}(\hat{x})$, and

$$D^* A D \hat{x} = \lambda \hat{x}, \quad \|\hat{x}\|_2 = 1;$$

and $\mathscr{S}$ is the eigenspace of A associated with the eigenvalues of A closest to $\hat{\lambda}$.

Corollary 4.3 (Eisenstat and Ipsen [11, Theorem 2.1]). *If A is Hermitian then*

$$\|\sin \Theta\|_2 \leqslant \|r\|_2 / \min_{\lambda \in \Lambda} |\lambda - \hat{\lambda}|.$$

If, in addition, D is nonsingular then

$$\|\sin \Theta\|_2 \leqslant \min\{\alpha_1, \alpha_2\} / \min_{\lambda \in \Lambda} \frac{|\lambda - \hat{\lambda}|}{|\hat{\lambda}|} + \|I - D\|_2,$$

where

$$\alpha_1 \equiv \|D^{-*} - D\|_2, \quad \alpha_2 \equiv \|I - D^{-*} D^{-1}\|_2.$$

The relative bound consists of two summands. The first summand represents the (absolute or relative) deviation of the equivalence transformation from a similarity, amplified by the relative eigenvalue separation; while the second summand represents the (absolute and relative) deviation of the similarity transformation from the identity. Hence the eigenspace $\mathscr{S}$ is insensitive to perturbations that are equivalence transformations if the equivalence transformation is close to a similarity transformation that does not differ much from the identity.

In the special case when $\mathscr{S}$ has dimension one, Corollary 4.3 is slightly stronger than [12, Theorem 2.2] and [11, Corollary 2.1].

Corollary 4.3 can be extended to bound angles between two eigenspaces of equal dimension $k \geqslant 1$, however at the expense of an additional factor $\sqrt{k}$ in the bound [11, Theorem 3.1]. The following bounds for equally dimensioned subspaces do without this factor.

Let $\mathscr{S}$ be an invariant subspace of A, and $\hat{\mathscr{S}}$ be an invariant subspace of D^*AD, where D is nonsingular and $\hat{\mathscr{S}}$ has the same dimension as $\mathscr{S}$. Let the columns of $\hat{X}$ be an orthonormal basis for $\hat{\mathscr{S}}$ and the columns of Y be an orthonormal basis for $\mathscr{S}^\perp$. Then

$$Y^*A = \Lambda Y^*, \quad D^*AD\hat{X} = \hat{X}\hat{\Lambda},$$

for some diagonal matrices Λ and $\hat{\Lambda}$. Below are Frobenius norm bounds on the angle between two equally dimensioned eigenspaces of a Hermitian matrix.

Theorem 4.4 (Li [21, Theorem 3.1]).[2] *If A is Hermitian and D is nonsingular then*

$$\|\sin \Theta\|_{\mathrm{F}} \leqslant \|(D - D^{-*})\hat{X}\|_{\mathrm{F}} / \min_{\lambda \in \Lambda, \hat{\lambda} \in \hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{|\lambda|} + \|(I - D^{-*})\hat{X}\|_{\mathrm{F}}$$

and

$$\|\sin \Theta\|_{\mathrm{F}} \leqslant \sqrt{\|(I - D)\hat{X}\|_{\mathrm{F}}^2 + \|(I - D^{-*})\hat{X}\|_{\mathrm{F}}^2} / \min_{\lambda \in \Lambda, \hat{\lambda} \in \hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{\sqrt{\lambda^2 + \hat{\lambda}^2}}.$$

These bounds give the same qualitative information as Corollary 4.3: The eigenspace is insensitive to perturbations that are congruence transformations if the congruence transformation is close to a similarity transformation that does not differ much from the identity. The first bound has the same form as the relative bound in Corollary 4.3. In particular, when $\mathscr{S}$ and $\hat{\mathscr{S}}$ have dimension one, the first bound in Theorem 4.4 implies

$$\|\sin \Theta\|_2 \leqslant \|(D - D^{-*})\|_2 / \min_{\lambda \in \Lambda, \hat{\lambda} \in \hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{|\lambda|} + \|I - D^{-*}\|_2,$$

which is almost identical to the relative bound in Corollary 4.3. Note that the relative eigenvalue separation in the second bound is different. More generally, Theorem 4.4 holds in any unitarily invariant norm [21, Theorem 3.2].

Multiplicative eigenvector bounds for the hyperbolic eigenvalue problem $Ax = \lambda Jx$, where A is Hermitian positive-definite and J is a diagonal matrix with unit diagonal entries are given in [31, Section 4].

[2] Here we have exchanged the roles of A and D^*AD compared to Theorem 3.1 in [21].

References

[1] J. Barlow, J. Demmel, Computing accurate eigensystems of scaled diagonally dominant matrices, SIAM J. Numer. Anal. 27 (1990) 762–791.

[2] J. Barlow, I. Slapničar, Optimal perturbation bounds for the Hermitian eigenvalue problem, Technical Report 99-001, Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA, February 1999.

[3] J. Barlow, P. Yoon, H. Zha, An algorithm and a stability theory for downdating the ULV decomposition, BIT 36 (1996) 14–40.

[4] C. Davis, W. Kahan, Some new bounds on perturbation of subspaces, Bull. Amer. Math. Soc. 75 (1969) 863–868.

[5] C. Davis, W. Kahan, The rotation of eigenvectors by a perturbation, III, SIAM J. Numer. Anal. 7 (1970) 1–46.

[6] P. Deift, J. Demmel, L. Li, C. Tomei, The bidiagonal singular value decomposition and Hamiltonian mechanics, SIAM J. Numer. Anal. 28 (1991) 1463–1516.

[7] J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselić, Z. Drmač, Computing the singular value decomposition with high relative accuracy, Linear Algebra Appl. 299 (1999) 21–80.

[8] J. Demmel, K. Veselić, Jacobi's method is more accurate than QR, SIAM J. Matrix Anal. Appl. 13 (1992) 1204–1245.

[9] I. Dhillon, A New $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem, Ph.D. thesis, Computer Science, University of California, Berkeley, CA, May 1997.

[10] I. Dhillon, G. Fann, B. Parlett, in: M. Heath et al. (Eds.), Application of a new algorithm for the symmetric eigenproblem to computational quantum chemistry, in: Proceedings of the Eighth SIAM Conference on Parallel Processing for Scientific Computing, SIAM, Philadelphia, 1997.

[11] S. Eisenstat, I. Ipsen, Relative perturbation bounds for eigenspaces and singular vector subspaces, in Applied Linear Algebra, SIAM, Philadelphia, 1994, pp. 62–65.

[12] S. Eisenstat, I. Ipsen, Relative perturbation techniques for singular value problems, SIAM J. Numer. Anal. 32 (1995) 1972–1988.

[13] S. Eisenstat, I. Ipsen, Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices, BIT 38 (1998) 502–509.

[14] S. Eisenstat, I. Ipsen, Three absolute perturbation bounds for matrix eigenvalues imply relative bounds, SIAM J. Matrix Anal. Appl. 20 (1998) 149–158.

[15] I. Gohberg, R. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, Wiley, New York, 1986.

[16] G. Golub, C. van Loan, Matrix Computations, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.

[17] M. Gu, S. Eisenstat, Relative perturbation theory for eigenproblems, Research Report YALEU/DCS/RR-934, Department of Computer Science, Yale University, 1993.

[18] I. Ipsen, Absolute and relative perturbation bounds for invariant subspaces of matrices, Linear Algebra Appl. 309 (2000) 45–56.

[19] I. Ipsen, Relative perturbation results for matrix eigenvalues and singular values, in: Acta Numerica 1998, Vol. 7, Cambridge University Press, Cambridge, 1998, pp. 151–201.

[20] R. Li, Relative perturbation theory: I. eigenvalue variations, SIAM J. Matrix Anal. Appl. 19 (1998) 956–982.

[21] R. Li, Relative perturbation theory: II. eigenspace and singular subspace variations, SIAM J. Matrix Anal. Appl. 20 (1999) 471–492.

[22] T. Londrè, N. Rhee, A note on relative perturbation bounds, SIAM J. Matrix Anal. Appl. 21 (2000) 357–361.

[23] R. Mathias, Fast accurate eigenvalue methods for graded positive-definite matrices, Numer. Math. 74 (1996) 85–103.

[24] R. Mathias, A bound for matrix square root with application to eigenvector perturbation, SIAM J. Matrix Anal. Appl. 18 (1997) 861–867.

[25] R. Mathias, Spectral perturbation bounds for positive definite matrices, SIAM J. Matrix Anal. Appl. 18 (1997) 959–980.

[26] R. Mathias, K. Veselić, A relative perturbation bound for positive definite matrices, Linear Algebra Appl. 270 (1998) 315–321.

[27] B. Parlett, Invariant subspaces for tightly clustered eigenvalues of tridiagonals, BIT 36 (1996) 542–562.

[28] B. Parlett, I. Dhillon, Relatively robust representations of symmetric tridiagonals, Department of Mathematics, University of California, Berkeley, February 1999.

[29] I. Slapničar, Accurate computation of singular values and eigenvalues of symmetric matrices, Math. Commun. 1 (1996) 153–167.
[30] I. Slapničar, N. Truhar, Relative perturbation bound for invariant subspaces of graded indefinite matrices, Technical Report, University of Split and University of Osijek, Croatia, September 1998.
[31] I. Slapničar, N. Truhar, Relative Perturbation theory for hyperbolic eigenvalue problem, tech. rep., University of Split and University of Osijek, Croatia, September 1998.
[32] I. Slapničar, K. Veselić, Perturbations of the eigenprojections of a factorised Hermitian matrix, Linear Algebra Appl. 218 (1995) 273–280.
[33] I. Slapničar, K. Veselić, A bound for the condition of a hyperbolic eigenvector matrix, Linear Algebra Appl. 290 (1999) 247–255.
[34] G. Stewart, J. Sun, Matrix Perturbation Theory, Academic Press, San Diego, 1990.
[35] N. Truhar, I. Slapničar, Relative perturbation of invariant subspaces, Math. Commun. 1 (1996) 169–174.
[36] N. Truhar, I. Slapničar, Relative perturbation bound for invariant subspaces of graded indefinite Hermitian matrices, Linear Algebra Appl. 301 (1999) 171–185.
[37] K. Veselić, I. Slapničar, Floating-point perturbations of Hermitian matrices, Linear Algebra Appl. 195 (1993) 81–116.